

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: DUAN, ROXANNE
RUBEN, STEVEN

(ii) TITLE OF INVENTION: Parotid Secretory Protein

(iii) NUMBER OF SEQUENCES: 18

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: HUMAN GENOME SCIENCES, INC.
(B) STREET: 9410 KEY WEST AVENUE
(C) CITY: ROCKVILLE
(D) STATE: MD
(E) COUNTRY: US
(F) ZIP: 20850

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: BROOKES, ANDERS A.
(B) REGISTRATION NUMBER: 36,373
(C) REFERENCE/DOCKET NUMBER: PF348

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (301) 309-8504
(B) TELEFAX: (301) 301-8439

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1028 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 49..795

(ix) FEATURE:

(A) NAME/KEY: sig_peptide
(B) LOCATION: 49..100

(ix) FEATURE:

(A) NAME/KEY: mat_peptide
(B) LOCATION: 103..795

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CACGAGATT CATGAGCATC CTCCTCTAAA CGCGTGTCAA GACAAAAG ATG CTT CAG
Met Leu Gln

5 CTT TGG AAA CTT GTT CTC CTG TGC GGC GTG CTC ACT GGG ACC TCA GAG 105
 Leu Trp Lys Leu Val Leu Leu Cys Gly Val Leu Thr Gly Thr Ser Glu
 -15 -10 -5 1

10 TCT CTT CTT GAC AAT CTT GGC AAT GAC CTA AGC AAT GTC GTG GAT AAG 153
 Ser Leu Leu Asp Asn Leu Gly Asn Asp Leu Ser Asn Val Val Asp Lys
 5 10 15

15 CTG GAA CCT GTT CTT CAC GAG GGA CTT GAG ACA GTT GAC AAT ACT CTT 201
 Leu Glu Pro Val Leu His Glu Gly Leu Glu Thr Val Asp Asn Thr Leu
 20 25 30

20 AAA GGC ATC CTT GAG AAA CTG AAG GTC GAC CTA GGA GTG CTT CAG AAA 249
 Lys Gly Ile Leu Glu Lys Leu Lys Val Asp Leu Gly Val Leu Gln Lys
 35 40 45

25 TCC AGT GCT TGG CAA CTG GCC AAG CAG AAG GCC CAG GAA GCT GAG AAA 297
 Ser Ser Ala Trp Gln Leu Ala Lys Gln Lys Ala Gln Glu Ala Glu Lys
 50 55 60 65

30 TTG CTG AAC AAT GTG ATT TCT AAG CTG CTT CCA ACT AAC ACG GAC ATT 345
 Leu Leu Asn Asn Val Ile Ser Lys Leu Leu Pro Thr Asn Thr Asp Ile
 70 75 80

35 TTT GGG TTG AAA ATC AGC AAC TCC CTC ATC CTG GAT GTC AAA GCT GAA 393
 Phe Gly Leu Lys Ile Ser Asn Ser Leu Ile Leu Asp Val Lys Ala Glu
 85 90 95

40 CCG ATC GAT GAT GGC AAA GGC CTT AAC CTG AGC TTC CCT GTC ACC GCG 441
 Pro Ile Asp Asp Gly Lys Gly Leu Asn Leu Ser Phe Pro Val Thr Ala
 100 105 110

45 AAT GTC ACT GTG GCC GGG CCC ATC ATT GGC CAG ATT ATC AAC CTG AAA 489
 Asn Val Thr Val Ala Gly Pro Ile Ile Gly Gln Ile Ile Asn Leu Lys
 115 120 125

50 GCC TCC TTG GAC CTC CTG ACC GCA GTC ACA ATT GAA ACT GAT CCC CAG 537
 Ala Ser Leu Asp Leu Leu Thr Ala Val Thr Ile Glu Thr Asp Pro Gln
 130 135 140 145

55 ACA CAC CAG CCT GTT GCC GTC CTG GGA GAA TGC GCC AGT GAC CCA ACC 585
 Thr His Gln Pro Val Ala Val Leu Gly Glu Cys Ala Ser Asp Pro Thr
 150 155 160

60 AGC ATC TCA CTT TCC TTG CTG GAC AAA CAC AGC CAA ATC ATC AAC AAG 633
 Ser Ile Ser Leu Ser Leu Leu Asp Lys His Ser Gln Ile Ile Asn Lys
 165 170 175

65 TTC GTG AAT AGC GTG ATC AAC ACG CTG AAA AGC ACT GTA TCC TCC CTG 681
 Phe Val Asn Ser Val Ile Asn Thr Leu Lys Ser Thr Val Ser Ser Leu
 180 185 190

70 CTG CAG AAG GAG ATA TGT CCA CTG ATC CGC ATC TTC ATC CAC TCC CTG 729
 Leu Gln Lys Glu Ile Cys Pro Leu Ile Arg Ile Phe Ile His Ser Leu
 195 200 205

75 GAT GTG AAT GTC ATT CAG CAG GTC GTC GAT AAT CCT CAG CAC AAA ACC 777
 Asp Val Asn Val Ile Gln Gln Val Val Asp Asn Pro Gln His Lys Thr
 210 215 220 225

80 CAG CTG CAA ACC CTC ATT TGAAGAGGAC GAATGAGGAG GACCACTGTG 825
 Gln Leu Gln Thr Leu Ile
 230

85 GTGCATGCTG ATTGGTTCCC AGTGGCTTGC CCCACCCCT TATAGCATCT CCCTCCAGGA 885

90 AGCTGCTGCC ACCACCTAAC CAGCGTGAAA GCCTGAGTCC CACCAGAAGG ACCTTCCCAG 945

10020339.121801

ATACCCCTTC TCCTCACAGT CAGAACAGCA GCCTCTACAC ATGTTGTCCT GCCCCTGGCA 1005
ATAAAGGCCC ATTTCTGCAA AAA 1028

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 249 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Leu Gln Leu Trp Lys Leu Val Leu Leu Cys Gly Val Leu Thr Gly
-18 -15 -10 -5
Thr Ser Glu Ser Leu Leu Asp Asn Leu Gly Asn Asp Leu Ser Asn Val
1 5 10
Val Asp Lys Leu Glu Pro Val Leu His Glu Gly Leu Glu Thr Val Asp
15 20 25 30
Asn Thr Leu Lys Gly Ile Leu Glu Lys Leu Lys Val Asp Leu Gly Val
35 40 45
Leu Gln Lys Ser Ser Ala Trp Gln Leu Ala Lys Gln Lys Ala Gln Glu
50 55 60
Ala Glu Lys Leu Leu Asn Asn Val Ile Ser Lys Leu Leu Pro Thr Asn
65 70 75
Thr Asp Ile Phe Gly Leu Lys Ile Ser Asn Ser Leu Ile Leu Asp Val
80 85 90
Lys Ala Gly Pro Ile Asp Asp Gly Lys Gly Leu Asn Leu Ser Phe Pro
95 100 105 110
Val Thr Ala Asn Val Thr Val Ala Gly Pro Ile Ile Gly Gln Ile Ile
115 120 125
Asn Leu Lys Ala Ser Leu Asp Leu Leu Thr Ala Val Thr Ile Glu Thr
130 135 140
Asp Pro Gln Thr His Gln Pro Val Ala Val Leu Gly Glu Cys Ala Ser
145 150 155
Asp Pro Thr Ser Ile Ser Leu Ser Leu Leu Asp Lys His Ser Gln Ile
160 165 170
Ile Asn Lys Phe Val Asn Ser Val Ile Asn Thr Leu Lys Ser Thr Val
175 180 185 190
Ser Ser Leu Leu Gln Lys Glu Ile Cys Pro Leu Ile Arg Ile Phe Ile
195 200 205
His Ser Leu Asp Val Asn Val Ile Gln Gln Val Val Asp Asn Pro Gln
210 215 220
His Lys Thr Gln Leu Gln Thr Leu Ile
225 230

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 235 amino acids
(B) TYPE: amino acid

10020339-121601

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met. Phe Gln Leu Gly Ser Leu Val Val Leu Cys Gly Leu Leu Ile Gly
1 5 10 15
Asn Ser Glu Ser Leu Leu Gly Glu Leu Gly Ser Ala Val Asn Asn Leu
20 25 30
Lys Ile Leu Asn Pro Pro Ser Glu Ala Val Pro Gln Asn Leu Asn Leu
35 40 45
Asp Val Glu Leu Leu Gln Gln Ala Thr Ser Trp Pro Leu Ala Lys Asn
50 55 60
Ser Ile Leu Glu Thr Leu Asn Thr Ala Asp Leu Gly Asn Leu Lys Ser
65 70 75 80
Phe Thr Ser Leu Asn Gly Leu Leu Leu Lys Ile Asn Asn Leu Lys Val
85 90 95
Leu Asp Phe Gln Ala Lys Leu Ser Ser Asn Gly Asn Gly Ile Asp Leu
100 105 110
Thr Val Pro Leu Ala Gly Glu Ala Ser Leu Val Leu Pro Phe Ile Gly
115 120 125
Lys Thr Val Asp Ile Ser Val Ser Leu Asp Leu Ile Asn Ser Leu Ser
130 135 140
Ile Lys Thr Asn Ala Gln Thr Gly Leu Pro Glu Val Thr Ile Gly Lys
145 150 155 160
Cys Ser Ser Asn Thr Asp Lys Ile Ser Ile Ser Leu Leu Gly Arg Arg
165 170 175
Leu Pro Ile Ile Asn Ser Ile Leu Asp Gly Val Ser Thr Leu Leu Thr
180 185 190
Ser Thr Leu Ser Thr Val Leu Gln Asn Phe Leu Cys Pro Leu Leu Gln
195 200 205
Tyr Val Leu Ser Thr Leu Asn Pro Ser Val Leu Gln Gly Leu Leu Ser
210 215 220
Asn Leu Leu Ala Gly Gln Val Gln Leu Ala Leu
225 230 235

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 235 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Phe Gln Leu Gly Ser Leu Val Val Leu Cys Gly Leu Leu Ile Gly
1 5 10 15

5 Thr Ser Glu Ser Leu Leu Gly Asp Val Ala Asn Ala Val Asn Asn Leu
20 25 30

Asp Ile Leu Asn Ser Pro Ser Glu Ala Val Ala Gln Asn Leu Asn Leu
35 40 45

10 Asp Val Gly Ser Leu Gln Gln Ala Thr Thr Trp Pro Ser Ala Lys Asp
50 55 60

Ser Ile Leu Glu Thr Leu Asn Lys Val Glu Leu Gly Asn Ser Asn Gly
65 70 75 80

15 Phe Thr Pro Leu Asn Gly Leu Leu Leu Arg Val Asn Lys Phe Arg Val
85 90 95

20 Leu Asp Leu Gln Ala Gly Leu Ser Ser Asn Gly Lys Asp Ile Asp Leu
100 105 110

Lys Leu Pro Leu Val Phe Glu Ile Ser Phe Ser Leu Pro Val Ile Gly
115 120 125

25 Pro Thr Leu Asp Val Ala Val Ser Leu Asp Leu Leu Asn Ser Val Ser
130 135 140

Val Gln Thr Asn Ala Gln Thr Gly Leu Pro Gly Val Thr Leu Gly Lys
145 150 155 160

30 Cys Ser Gly Asn Thr Asp Lys Ile Ser Ile Ser Leu Leu Gly Arg Arg
165 170 175

35 Leu Pro Phe Val Asn Arg Ile Leu Asp Gly Val Ser Gly Leu Leu Thr
180 185 190

Gly Ala Val Ser Ile Leu Leu Gln Asn Ile Leu Cys Pro Val Leu Gln
195 200 205

40 Tyr Leu Leu Ser Thr Met Ser Gly Ser Ala Ile Gln Gly Leu Leu Ser
210 215 220

Asn Val Leu Thr Gly Gln Leu Ala Val Pro Leu
225 230 235

45

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
50 (A) LENGTH: 206 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear.

55 (ii) MOLECULE TYPE: protein

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Phe Gln Leu Gly Ser Leu Val Val Leu Cys Gly Leu Leu Ile Gly
1 5 10 15

65 Thr Ser Gly Ser Leu Phe Asp Ile Phe Gln Asn Pro Glu Leu Asp Val
20 25 30

Glu Ser Val Trp Ser Glu Ile Asn Tyr Arg Ile Arg Tyr Ala Leu Glu
35 40 45

10020139-121801

Thr Met Asp Leu Asp Met Leu Ala Asp Tyr Leu Ser Lys Arg Gly Ile
 50 55 60
 5 Glu Leu Lys Ile Lys Asp Leu Arg Ile Leu Asn Leu Asn His Glu Val
 65 70 75 80
 Ser Pro Asn Gly Asp Glu Val Thr Leu Lys Met Pro Met Ala Leu Asn
 85 90 95
 10 Ala Ser Leu Ser Leu Pro Ala Arg Asp Leu Thr Thr Asp Val Ser Ile
 100 105 110
 Ser Met Glu Ala Ile Thr Ser Phe Ala Ile Glu Lys Asp Pro Lys Thr
 115 120 125
 15 Gly Arg Arg Val Leu Asn Met Gln Arg Cys Ser Leu Asn Thr Asp Asn
 130 135 140
 Thr Ser Ile Ser Leu Leu Asn Arg Lys Ser Asn Phe Val Asn Leu Ala
 145 150 155 160
 Leu Asp Ser Ala Leu Tyr Leu Ile Lys Arg Gly Leu Thr Leu Pro Val
 165 170 175
 25 Arg Arg Gln Leu Cys Pro Val Leu Gln Leu Ile Ile Ser Asn Thr Phe
 180 185 190
 His Pro Asp Glu Ile Ser Asn Pro Gln Thr Ala Ile Ser Thr
 195 200 205

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CTACAGCCAT GGAGTCTCTT CTGACAATC TTGGCAATG

39

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CATCGCGGAT CCAATGAGGG TTGCGAGCTG GGTTTT

36

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 43 base pairs
 (B) TYPE: nucleic acid

10000134-121301

5

10

43

15

20

25

31

30

35

40

45

60

120

180

240

.300

360

420

449

60

65

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

5 ATGAGCATCC TCCTCTAAAC GCGTGTCAAG ACAAAGATG CTNNCAGCTT TGGAACTTG 60
 TTCTCCTGTG CGGCGTGCTC ACTGGGACCT CAGAGTCTCT TCTTGACAAT CTGGCAATG 120
 10 ACCTAAGCAA TGTCGTGGAT AAGCTGGAAC CTGTTCTTCA CGAGGGACTT GAGACAGTTG 180
 ACAATACTCT TAAAGGCATC CTTGAGAAAC TGAAGGTCGA CCTAGGAGTG CTTCAGAAAT 240
 CCAGTGCTTG GCAACTGGCC AACAGAAGN CCAGGAAGCT GAGAAATTGC TGAACAATGT 300
 15 CATTCTAAG CTGCTTCCAA CTAACACGGA CATTTTGGG TTGAAANTCA GCAATNCCN 360
 CANCCGGATG TTCAAAGNTG NANCATCGA TGATGGGCAA AGGCTTTAAN CCGGAGGCTT 420
 20 CCCTGTTTAC CGGAATGTT CAANGTNGC CCGGGCCNT CATTGGGCCA GNTTATCAAA 480
 NCTGGAAAGC TTCCTGGGAC CTCCGGACNG GNTCAACAAT TGAAANGATT CCCCANA 538

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 359 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic).

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGCAGGAGAT TTCATGAGCA TCCTCTCTA AACGCGTGC AAGACAAAAG ATGCTTCAGC 60
 40 TTTGGAACT TGTCTCCTG TCGCGGTGC TCACTGGGAC CTCAGAGTCT CTTCTTGACA 120
 ATCTTGCAA TGACCTAAGC AATGTCGTGG ATAAGCTGGG AACCTGTTCT TCACGAGGA 180
 45 CTTGAGACAG TTGACAATAC TCTTAAAGGC ATCCTTGAGA AACTGAAGGT CGANCTAGGA 240
 GTGCTTCAGA AATCCAGTGC TTGGCAACTG GNCAAGCAGA AGGNCCAGG AAGCTGAGAA 300
 50 ATTGCTGGAN CAATGTCAAT TCTAAGCTGN TTCCGACTAA CACGGNCATT TTTGGGTTG 359

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 374 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: DNA (genomic).

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

65 ACCCAGCGT CCGGCGTGAT CAACACGCTG AAAAGCACTG TATCCTCCCT GCTGCAGAAG 60
 GAGATATGTC CACTGATCCG CATCTTCATC CACTCCCTGG ATGTGAATGT CATTGAGCAG 120
 GTCGTCGATA ATCCTCAGCA CAAAACCCAG CTGCAACCC TCATCTGAAG AGGACGAATG 180

1002019-121604

AGGAGGACCA CTGTGGTGCA TGCTGATTGG TTCCAGTGG CTTGCCCCAC CCCCTTATAG 240
 CATCTCCCTC CAGGAAGCTG CTGCCACCAC CTAACCAGCG TGAAAGCCTG GAGTCCCACC 300
 AGAAGGACCT TCCCAGATAC CCCTTTTTC TCACAGTCAG AGNGGNNGC CTCTTACACN 360
 TGTGTCCNG GCCC 374

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 406 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTCATGAGCA TCCTCCTCTA AACGCGTGTC AAGACAAAAG ATGCTTCACT TTGGAAACTT 60
 GTTCTCCTGT GCGGCGTGCT CACTGGGACC TCAGAGTCTC TTCTTGACAA TCTTGGAAT 120
 GACCTAAGCA ATGTCGTGGA TAAGCTGGAA CCTGTTCTTC ACGAGGGACT TGAGACAGTT 180
 GACAATACTC TTAAAGGCAT CCTTGAGAAA CTGAAGGTCG ACCTAGGAGT GCTTCAGAAA 240
 TCCAGTGCTT GGCAACTGGG CCANCAGAAA GGCCAGGGA AAGCGGAGAA ATTGCTGGAA 300
 CAATGTTTCA TTCTAAAGCT GCTTTCCAAC TAACACGGGA CNTTTTTGGG GTTTGNAAAA 360
 TCAGCCAAC TCCCTCAACC NNGGATGTTT CAAAGCTGGA AACCGN 406

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 493 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

TCGACCCACG CGTCCGCCGC CCCCAACTTT TTGGCAGTTC TCCCATCTCT TGCCCACTTG 60
 ACCAGACTTT AATAGTTCCC TGTGTTTTCC AGACACAGCC AAATCATCAA CAAGTTCGTG 120
 AATAGCGTGA TCAACAGCCT GAAAAGCACT GTATCCTCCC TGCTGCAGAA GGAGATATGT 180
 CCACTGATCC GCATCTTCAT CCACTCCCTG GATGTGAATG TCATTCAGCA GGTCGTCGAT 240
 AATCCTCAGC ACAAACCCCA GCTGCAAACC CTCATCTGAA GAGGACGAAT GAGGAGGACC 300
 ACTGTGGTGC ATGCTGGTGA GGAGCCAGTC TCTGTGCCCC AATGCACAGG GGCCTATGGT 360
 GAAGTAAAG TCAAGCGTGG CTTCCTTAT TTTTGTGTTA GAAGACTGTG CCTTCATCTC 420
 AGTCATAGAT TGAGCCCTGG NCCCCATCCC ANGCTAAGGC CTGATTCTGG TCANACTCTG 480

10020139-11901

493

5

(i). SEQUENCE CHARACTERISTICS:

- 10

15

30

35

40

45

50

55

GTGGNCCGGG	GCCCATCATT	GGGCCAGANT	TATCAACCTG	AAAGCCTCCN	NGGANCTCCT	60
GACCGCAGTC	AACAATTGGA	AACTGGATCC	CCCAGAACAA	CAACCAGCCT	GGTTGNCCGT	120
NCTGGGGAAG	AATGCCGNCC	AATGAANCCC	AAACCAAGCA	NCTTCACTNN	TNCCTNGGCT	180
GGGGACCAAA	ACACCAGGCC	AAAATCCANT	NAANCAAGTN	TCCGTGGNAA	TAAGCGTGGA	240
ATCCAAACAA	CGCTGGGAAA	AAGCANTGGG	NATNCCNTCC	CTGGCTGGGC	AAGAAAGGGN	300
GATATGGTCC	ACTGGAATCC	GGAATTTTTA	ANCCAATTCC	CTNGGAATGT	GGNAATGTCA	360